

SUPPLEMENTARY DATA

We constructed dynamic models consisting of systems of difference equations. The U.S. population ages 0 to 19 years is modeled at one year intervals starting at year $t = 0$ (2001) and ending at year $t = 49$ (2050). Specifically, we define numbers of people in various states, rates, and flows. All rates are annual and flows occur during year t , i.e., in the interval $(t-1, t]$. Moreover, the notation and model suppresses the ten race-sex combinations, i.e., there are ten separate models described by the following. First we characterize the Census population projections.

$N(a, t)$ = the total number of people age $= a$ at year t coming from Census projections.

$b(t)$ = Census projection of the number of births in year t .

$m(a, t)$ = Census projection of the number of people age $= a$ migrating into the U.S. in year t .

$d(a-1, t)$ = Census projected death rate among all people age $= a-1$ in year t .

$D(a, t)$ = the number of deaths in year t among people age $= a-1$ at the beginning of the year. Note that

$$D(a, t) = d(a-1, t)N(a-1, t-1).$$

Using the above notation

$$N(0, t) = b(t) + m(0, t)$$

$$N(a, t) = N(a-1, t-1) - D(a, t) + m(a, t)$$

with $a = 1, 2, \dots, 19$ and $t = 1, \dots, 49$. For the diabetes projections let

$X(a, t)$ = the number of people age $= a$ without diabetes at year t .

$Y(a, t)$ = the number of people age $= a$ with Type 1 diabetes at year t .

$Z(a, t)$ = the number of people age $= a$ with Type 2 diabetes at year t .

$\delta(a-1, t)$ = non-diabetes death rate among $X(a-1, t-1)$ in year t .

$\delta(a-1, t)$ = diabetes death rate among $Y(a-1, t-1)$ in year t ; $r_2 \delta(a-1, t)$ = diabetes death rate among $Z(a-1, t-1)$ in year t ; r_1 and r_2 are relative risks which are assumed to be age and time invariant.

$\beta(a-1, t)$ = incidence rate of Type 1 diabetes among $X(a-1, t-1)$ in year t , estimated from SEARCH data as described in the previous section.

$\gamma(a-1, t)$ = incidence rate of Type 2 diabetes among $X(a-1, t-1)$ in year t , estimated from SEARCH data as described in the previous section.

$f_x(t)$ = proportion of $b(t)$ without diabetes.

$f_y(t)$ = proportion of $b(t)$ with Type 1 diabetes.

$f_z(t)$ = proportion of $b(t)$ with Type 2 diabetes.

$g_x(a, t)$ = proportion of $m(a, t)$ without diabetes.

$g_y(a, t)$ = proportion of $m(a, t)$ with Type 1 diabetes.

$g_z(a, t)$ = proportion of $m(a, t)$ with Type 2 diabetes.

Define $\theta(a, t) = \frac{Y(a, t) + Z(a, t)}{X(a, t) + Y(a, t) + Z(a, t)}$, the prevalence of diabetes among people age $= a$ at year t ;

$\theta_1(a, t) = \frac{Y(a, t)}{X(a, t) + Y(a, t) + Z(a, t)}$, the prevalence of Type 1 diabetes among people age $= a$ at year t ;

$\theta_2(a, t) = \frac{Z(a, t)}{X(a, t) + Y(a, t) + Z(a, t)}$, the prevalence of Type 2 diabetes among people age $= a$ at year

t , noting that $\theta(a, t) = \theta_1(a, t) + \theta_2(a, t)$. We impose the constraints $X(a, t) + Y(a, t) + Z(a, t) = N(a, t)$ so that the diabetes projections are consistent with the Census population projections.

To this end, we first set the initial conditions

$$X(a, 0) = [1 - p_1(a) - p_2(a)]N(a, 0)$$

$$Y(a, 0) = p_1(a)N(a, 0)$$

$$Z(a, 0) = p_2(a)N(a, 0)$$

SUPPLEMENTARY DATA

$a = 0, 1, \dots, 19$, with $p_1(a)$ and $p_2(a)$ the 2001 prevalences of Type 1 and Type 2 diabetes estimated from the SEARCH data as described in the second section and $N(a, 0)$ is the 2001 population from the Census data. We also need the relations $f_x(t) + f_z(t) + f_y(t) = 1$ and $g_x(a, t) + g_y(a, t) + g_z(a, t) = 1$ to guarantee that incoming flows match. Lastly, we need the deaths in the diabetes model to equal deaths in the Census projections model, or

$$\begin{aligned} & \delta(a-1, t)X(a-1, t-1) + \\ & r_1\delta(a-1, t)Y(a-1, t-1) + \\ & r_2\delta(a-1, t)Z(a-1, t-1) \\ & = d(a-1, t)N(a-1, t-1) \end{aligned}$$

which implies (with a little algebra)

$$\delta(a-1, t) = \frac{d(a-1, t)}{[1 - \theta(a-1, t-1)] + [r_1\theta_1(a-1, t-1) + r_2\theta_2(a-1, t-1)]}.$$

Consider the transition matrix

	$X(a, t):$	$Y(a, t):$	$Z(a, t):$	$D(a, t):$	
$X(a-1, t-1):$	$\alpha(a-1, t)$	$\beta(a-1, t)$	$\gamma(a-1, t)$	$\delta(a-1, t)$	
$Y(a-1, t-1):$	0	$[1 - r_1\delta(a-1, t)]$	0	$r_1\delta(a-1, t)$	(1)
$Z(a-1, t-1):$	0	0	$[1 - r_2\delta(a-1, t)]$	$r_2\delta(a-1, t)$	
$D(a-1, t-1):$	0	0	0	1	

where $a = 1, 2, \dots, 19$ and $t = 1, 2, \dots, 49$. Note that the rows of this matrix display the distribution of the beginning year stocks (age = $a-1$) to the ending year stocks (the columns, age = a), and thus the transition rates in each row must be nonnegative and add to unity for each year t . Also, some assumptions about transition rates are apparent. First, people cannot move from diabetes to non-diabetes. Second, people cannot move from one diabetes state to another diabetes state. Third, as stated above, the relative risks of death for the two diabetes states to the non-diabetes state are constant by age over time. The transition matrix (1) and the inflows lead to the system of difference equations

$$\begin{aligned} X(0, t) &= f_x(t)b(t) + g_x(0, t)m(0, t) \\ X(a, t) &= \alpha(a-1, t)X(a-1, t-1) + g_x(a, t)m(a, t) \\ Y(0, t) &= f_y(t)b(t) + g_y(0, t)m(0, t) \\ Y(a, t) &= \beta(a-1, t)X(a-1, t-1) + [1 - r_1\delta(a-1, t)]Y(a-1, t-1) \\ &\quad + g_y(a, t)m(a, t) \\ Z(0, t) &= f_z(t)b(t) + g_z(0, t)m(0, t) \\ Z(a, t) &= \gamma(a-1, t)X(a-1, t-1) + [1 - r_2\delta(a-1, t)]Z(a-1, t-1) \\ &\quad + g_z(a, t)m(a, t) \end{aligned} \tag{2}$$

where, again, $a = 1, 2, \dots, 19$ and $t = 1, 2, \dots, 49$. Initial conditions for (2) are $X(a, 0)$, $Y(a, 0)$, $Z(a, 0)$ as defined earlier. Finally, in all our projections we set $f_x(t) = 1$, $f_y(t) = 0$, $f_z(t) = 0$, or incoming births are non-diabetic, and

SUPPLEMENTARY DATA

$$g_x(a,t) = [1 - \theta_1(a,t-1) - \theta_2(a,t-1)],$$

$$g_y(a,t) = \theta_1(a,t-1),$$

$$g_z(a,t) = \theta_2(a,t-1)$$

All diabetes projections were calculated using the programming language GAUSS.¹

We finish this section with a matrix characterization of the system (2). Let $\mathbf{T}(a,t)$ be the 3x3 matrix of transition rates (excluding death rates) defined in the transition matrix (1) and set

$$\mathbf{T}(t) = \begin{bmatrix} \mathbf{O} & \mathbf{T}(0,t) & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{T}(1,t) & \dots & \mathbf{O} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} & \mathbf{T}(18,t) \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \end{bmatrix},$$

a 60x60 matrix of transitions with 3x3 zero matrices everywhere but the upper diagonal. If we define the 1x60 row vector of states

$$\mathbf{S}(t) = [X(0,t), Y(0,t), Z(0,t), \dots, X(19,t), Y(19,t), Z(19,t)],$$

and the 1x60 row vectors of flows

$$\mathbf{B}(t) = [f_x(t)b(t), f_y(t)b(t), f_z(t)b(t), 0, 0, 0, \dots, 0, 0, 0]$$

$$\mathbf{M}(t) = [g_x(0,t)m(0,t), g_y(0,t)m(0,t), g_z(0,t)m(0,t), \dots, \\ g_x(19,t)m(19,t), g_y(19,t)m(19,t), g_z(19,t)m(19,t)],$$

then the projections model can be represented in matrix form as

$$\mathbf{S}(t) = \mathbf{S}(t-1)\mathbf{T}(t) + \mathbf{B}(t) + \mathbf{M}(t)$$

with initial conditions $\mathbf{S}(0)$ as defined above.

Reference:

1. Aptech Systems Inc.. *GAUSS Mathematical and Statistical System version 10.0*. 2009 Maple Valley, WA: Aptech Systems Inc.